ABSTRACT

The purpose of this study was to determine the between-day reliability of commonly used strength measures in male youth athletes, while considering resistance training experience. Data were collected on 25 male athletes over two testing sessions, with 72 hours rest between, for the 3RM front squat, chin up and bench press. Subjects were initially categorized by resistance training experience (inexperienced; 6-12 months, experienced; >2 years). The assessment of the between-day reliability (coefficient of variation [CV%]) showed the front squat (experienced: 2.90%; inexperienced: 1.90%), chin up (experienced: 1.70%; inexperienced: 1.90%), and bench press (experienced: 4.50%; inexperienced: 2.40%) were all reliable measures of strength in both groups. Comparison between groups for the error of measurement for each exercise showed *trivial* differences. When both groups were combined, the CV% for the front squat, bench press, and chin up were 2.50%, 1.80%, and 3.70%, respectively. This study provides scientists and practitioners with the between-day reliability reference data to determine real and practical changes for strength in male youth athletes with different resistance training experience. Furthermore, this study demonstrates that 3RM front squat, chin up and bench press are reliable exercises to quantify strength in male youth athletes.

Key Words: Squat, Bench press, Chin up, Testing

INTRODUCTION

Muscular strength, is the ability to exert force on an external object or resistance, and is an important physical quality for athletic performance in youth (10) and senior (29) athletes. In sports that require high levels of strength (e.g. rugby union, rugby league, and American football) due to the contact nature of competition (20, 22), strength development (12, 28), and quantification (17, 27) are common. Strength testing adolescent athletes in the front squat (7, 33), chin up (5, 7), and bench press (3, 30) exercises are commonly used and enables comparisons to reference data for athletes of a similar age and sport (7, 18, 31). These data can also provide guidance of physical preparation (19) through the prescription of specific resistance loads (21). Furthermore, pre- and post- training intervention testing is often undertaken to assess the usefulness and efficacy of the prescribed training. However, to correctly assess whether an individual has improved their strength in a resistance training movement, it is necessary to determine whether a change is real or due to testing error. This can only be achieved when the between-day reliability of each movement is calculated and available (23, 26), thus should be a key consideration for the scientist and practitioner.

Athletes with different resistance training experience, determined as the length of time in months and years that an athlete has performed resistance training, may have varying levels of between-day reliability in strength measures (25). Previous work from Ritti-Dias et al. (25) explored this phenomenon, demonstrating that significant changes (p= 0.01) in strength occur between resistance training sessions in subjects with no prior resistance training experience while their well-trained counterparts (>2 years) do not. These differences were attributed to modification in motor unit recruitment and rate coding, which is more prone to occur if athletes have little to no prior resistance training experience (11). As such, scientists and practitioners should acknowledge this when interpreting strength changes in athletes with differing training experience. In contrast to this, Comfort and McMahon (4) reported extremely high reliability of both the power clean (ICC = 0.997) and back squat (ICC = 0.994) in inexperienced (resistance training history of <1 year), albeit senior athletes. As such, the between day reliability of specific populations and their relative training experience should be established to help practitioners and scientists make informed decisions regarding the effectiveness of a training intervention.

The between-day reliability of commonly used strength tests in youth athletes has yet to be reported, despite a plethora of research evaluating strength in this cohort (3, 7, 30-32). This potentially infers that it is unknown when a 'real' change in strength has been achieved beyond that of normal between-day variation. Hopkins (14) suggests that the use of the coefficient of variation (CV%), typical error (TE) and the smallest worthwhile change (SWC) should be used in conjunction to assess the sensitivity and usefulness of the exercise, and also to allow the scientist and practitioner to assess whether a change has been real (i.e. >TE) and of practical significance (i.e. >SWC). Previously, this method has established the between-day reliability of sprint ability (6) and fatigue responses (26) in adolescent rugby players, and is necessary for the accurate quantification of strength.

To this end, the purpose of this study was to establish the between-day reliability of the front squat, bench press, and neutral grip chin up in experienced and inexperienced resistance trained male athletes, who are aged between 16-18 years old.

Methods

Experimental Approach to the Problem

To assess the between-day reliability of experienced and inexperienced resistance trained adolescent rugby union players, players undertook strength assessments on two separate occasions at the same time of day separated by 72 hours. The strength tests included the front squat, bench press, and chin up, as these have commonly been used in adolescent resistance training literature (5, 7, 30). Reliability of the strength variables across the two testing sessions were assessed using TE and CV%, while differences in error between groups were assessed using magnitude based inferences (MBI's).

Subjects

Twenty-five male adolescent rugby athletes, aged 16-18 years, participated in the study (age 17.1 ± 0.3 years, height 178.3 ± 5.6 cm, and body mass 87.0 ± 10.7 kg). Fourteen subjects were *experienced* (age 17.3 ± 0.4 years, training age 2.7 ± 0.2 years, height 179.3 ± 5.8 cm, body mass 88.1 ± 12.2 kg) and eleven were *inexperienced* (n = 11, age 16.7 ± 0.3 years, training age 0.9 ± 0.2 years, height 177.4 ± 4.1 cm, body mass 85.7 ± 8.2 kg), according to their resistance training history. Experienced subjects had 2 or more years of continuous resistance training experience and inexperienced subjects had between 6-12 months of continuous resistance training experience, as previously categorised (4, 24, 25). Subjects with between 12 and 24 months' resistance training experience were not recruited to ensure discrete groups were created with clear differences in experience. All subjects were required to have had continuous resistance training history, 2-5 times per week, and were excluded if they had ceased resistance training for >1 month, due to injury or non-participation. All experimental procedures were approved by Leeds Beckett University's ethics committee and written assent was provided by all subjects following parental consent.

Procedures

All testing was conducted in the month of September, which is at the beginning of the adolescent rugby season, and was completed on two days separated by 72 hours. Resistance training in the prior two months had consisted of the same off-season conditioning programme for all participants. Lower- and upper- body strength were assessed using a 3-repetition maximum (3RM) strength test which included the front squat, neutral grip chin up, and bench press. The subjects in this study were all familiar with these movements and had previously used all movements during resistance training. In each session, subjects arrived and were measured for height and body mass at 10:00 hours. All equipment, environmental conditions (i.e. temperature), and footwear were unchanged across sessions. Weightlifting accessories (e.g., belts and wrist straps) were not used in either trial. Subjects were instructed to refrain from physical activity and maintain normal dietary habits for 48 hours prior to each testing sessions. The subjects were informed of the order of testing and completed a standardised warm up, which consisted of stationary cycling, dynamic movements and stretches prior to the initiation of any external resistance. Upon the completion of the warm up, an exercise specific warm up was completed that included 8 repetitions with an empty bar (or body weight for the chin up exercise), followed by two sets of 5 repetitions, and then 3 repetitions all at submaximal self-selected loads as previously completed in adolescent resistance training literature (7). Each subject had three attempts to achieve a 3RM load, with minimum incremental increases in load being 2.5 kg, and were required to have three minutes rest between maximal attempts (7, 32).

For the front squat, subjects were required to squat with a barbell (Eleiko Sport AB, Halmstad, Sweden) resting across the front of the shoulders, until the top of the thigh was parallel with the floor. Heels were to remain in contact with the ground throughout the movement, and the subject was to return to the initial standing position. The bench press required subjects to select a comfortable grip on the barbell (Eleiko Sport AB, Halmstad, Sweden) and were instructed to lower the bar to touch the chest and return to the starting point with the arms fully extended without any external assistance. The neutral grip chin up consisted of the subjects starting with their arms fully extended and pulling themselves to a position where the eyes were above the chinning bar. 3RM strength for the chin up was recorded as the participant's body mass plus external load, with external load being attached through a weighted belt (Harbinger, Leather Dip Belt, USA). This testing methodology has previously been used within the literature (7, 31, 32). All measures were assessed and approved by the lead researcher.

Statistical Analyses

Data are presented as either mean ± standard deviation (SD) or means with 90% confidence intervals (90% CI) where specified. Between-day reliability was assessed using a Microsoft Excel spreadsheet (14) which allowed for the SWC, CV%, TE, and usefulness of each test to be calculated. The usefulness of the test was assessed by comparing the SWC with the CV% (15). If the CV% was less than the SWC, the test was deemed to be "*good*". If the CV% was the same as the SWC then the test was considered as "*OK*", and if the CV% was greater than the SWC then the test was classed as "*poor*". A CV% of <5% was set as the criterion to declare that a variable was reliable, which has previously been used throughout the literature (1, 2, 6).

Magnitude based inferences (MBIs) were used to compare the between day TE for each group using a Microsoft Excel spreadsheet (16). If differences were evident (i.e. greater than the SWC), the CV% and TE for each group and exercise were the originally reported value. If differences were *trivial,* experience levels were combined and the TE and CV% were calculated for the combined groups. The probability that the magnitude of the difference were greater than the smallest worthwhile change were rated as <0.5%, *almost certainly not*; 0.5-5%, *very unlikely*; 5-25%, *unlikely*; 25-75%, *possibly*; 75-95%, *likely*; 95-99.5%, *very likely*; >99.5% *almost certainly*. Differences less than the SWC were described as *trivial*. In cases where the 90% CI crossed the lower and upper boundary of the SWC (ES±0.2), the magnitude of the difference was described as *unclear.*

RESULTS

Tables 1 and 2 present the TE, CV%, SWC, and usefulness rating of all exercises in the experienced and inexperienced resistance trained groups. All exercises were deemed to be reliable (i.e. CV% <5%) with *good* usefulness, apart from the bench press in the experienced group, which was classed as *poor*.

Insert Table 1 and 2 here

The error of the standardized TE between the two groups were *almost certainly trivial* for the front squat and chin up, and *likely trivial* for the bench press. Therefore, due to this *trivial* error, Table 3 outlines the TE and CV% of the front squat, chin up and bench press when the two experience levels are combined.

Insert Table 3 here

DISCUSSION

Given the importance of strength for athletic performance (29), and the lack of between day reliability data for commonly used strength measures in youth athletes, this study evaluated the between-day reliability of the front squat, chin up, and bench press in experienced and inexperienced resistance trained male youth athletes. Findings showed both groups had acceptable levels (i.e. CV% <5%) of reliability for the front squat, bench press and chin up, which suggests that these tests are reliable. This study also found the differences in error between experience levels were *trivial,* thus when both groups were combined all exercises still had acceptable levels of reliability. This suggests that between-day reliability by experience level may not be warranted in practice.

The high reliability of maximal testing found in this study agrees with previous research which has investigated the power clean and back squat exercises in experienced, inexperienced, and youth athletes (4, 9). However, this is the first study to analyse the differences in experienced and inexperienced adolescent athletes for the front squat, bench press and chin up. Current research suggests strength development following resistance training interventions in adolescent athletes range from small (7.6% in the bench press across a 12-week period; (12)) to very large (72.5% in the squat across a 15-week period (28)), dependent upon the exercises employed. These findings imply that the lower body exercise of the squat appears to improve to a greater extent than upper body strength exercises (i.e. the bench press and chin up) over similar lengths of time (5, 12, 28). While improvements in mean strength in the aforementioned studies have surpassed the CV% reported in the current study, this has not taken into account individual deviation and change.

Using the method proposed by Hopkins (14) outlining test rating, all measures excluding the bench press in experienced athletes were deemed to be *good* for determining actual and practical changes in strength with the CV% being smaller than the SWC. However, for athletes with over two years' resistance training experience, the bench press was rated as having poor usefulness due to the CV% of 4.50% being greater than the SWC of 3.51%. This suggests the test has poor sensitivity and that coaches of individuals who have >2 years' experience may be required to interpret scores of bench press with a greater level of prudence when assessing whether an actual improvement has occurred. Despite this, real change can still be identified with 75% certainty when change exceeds the sum of CV% and the cohort's SWC (6, 13, 26).

While the current study suggests that maximal strength testing at varying experience levels is reliable, previous differences found in the reliability of strength testing may solely come down to the definition of *inexperienced* and *experienced* athletes. It appears that when healthy young adults who have no previous resistance training experience complete strength testing, maximal strength between sessions may vary due to a lack of familiarization with the movement (25). However, as seen in the current study, 6 months resistance training experience appears to be a sufficient period of time to minimize any effects of familiarization and off-set the initial rapid adaptation of the central nervous system to external loading (8, 11, 25). This 6 month time frame has also been shown to be sufficient in promoting reliability in other resistance training exercises in young adults and adolescents (4, 9). Additionally, this familiarization to learning a movement appears to be permanent to an extent, with periods of abstinence (>6 months) from resistance training in experienced individuals not appearing to affect the reliability of their maximal lifting capability (11, 25).

This study, is not without its limitations. In the current study, depth in the squat was assessed by the top of the thigh being parallel to the floor (which has previously been used in adolescent literature (7)), however this could lead to systematic bias with interrater reliability, as this criterion is yet to be established, especially in between-day studies. Due to the design of the study, it was not possible to determine at what time point strength testing is reliable for an adolescent athlete so that maximal strength testing can be implemented, thus this may be a direction for future work. Finally, due to the variation of the CV% and TE which is specific to a population, similar numbers cannot be generalized to other exercises and individuals of differing ages. Moreover, this analysis is only specific to 16-18 year old male athletes that are completing the front squat, chin up and bench press. Athletes of a different sex and age may not follow similar patterns of reliability in these movements.

In conclusion, this study established that the 3RM front squat, chin up and bench press were reliable in experienced and inexperienced male youth athletes. Due to the *trivial* differences in error between the experienced and inexperienced resistance trained groups, the between-day reliability can be grouped together and still maintain a high level of reliability and precision. If experience levels are combined, practitioners and scientists can use a CV% of 2.50%, 1.80%, and 3.70% for the front squat, chin up, and bench press when determining change in male adolescent athletes. This study provides scientists and practitioners with between-day reliability reference data to determine real and practical changes for youth athletes with different resistance training experience. Additionally, this study also provides confirmation that improvements in strength, commonly seen in adolescent research (4, 11), are of a 'real' and practical nature.

PRACTICAL APPLICATIONS

The high level of reliability in both the experienced and inexperienced resistance trained groups allow a combined CV% of 2.50, 1.80 and 3.70, for the front squat, chin up, and bench press to be used, respectively. For the practitioner, an example of how this could be used is if an under 18 level male athlete managed to front squat 110kg, then they can be assured that this individual's front squat strength is within the range of 107.25 – 112.75kg. Furthermore, if the coach of an inexperienced resistance trained athlete wanted to assess whether a player had improved in the bench press, the coach could utilise both the SWC and CV%. An example of this would use the bench press and the corresponding SWC of 3.00% and CV% of 2.40%. If this athlete at the beginning of the intervention managed to bench press 75 kg, and at the end managed to bench press 80 kg, then the coach could assume that there had been a real change in strength as the improvement was greater than the SWC and error of the measurement. However, if the athlete had only managed to bench press 78kg at the post-test, then the coach would not be able to definitively state that an improvement had been made due to the error of the test crossing into the SWC.

References

1. Buchheit M, Lefebvre B, Laursen PB, and Ahmaidi S. Reliability, usefulness, and validity of the 30-15 Intermittent Ice Test in young elite ice hockey players. *J Strength Cond Res* 25: 1457-1464, 2011.
2. Buchheit M, Spencer M, and Ahmaidi S. Reliability, usefulness, and validity of a repeated sprint and jump ability test. *Int J Sports Physiol Perform* 5: 3-17, 2010.
3. Christou M, Smilios I, Sotiropoulos K, Volaklis K, Pilianidis T, and Tokmakidis SP. Effects of resistance training on the physical capacities of adolescent soccer players. *J Strength Cond Res* 20: 783-791, 2006.
4. Comfort P and McMahon JJ. Reliability of Maximal Back Squat and Power Clean Performances in Inexperienced Athletes. *J Strength Cond Res* 29: 3089-3096, 2015.
5. Coutts AJ, Murphy AJ, and Dascombe BJ. Effect of direct supervision of a strength coach on measures of muscular strength and power in young rugby league players. *J Strength Cond Res* 18: 316-323, 2004.
6. Darrall-Jones JD, Jones B, Roe G, and Till K. Reliability and Usefulness of Linear sprint testing in adolescent Rugby Union and League players. *J Strength Cond Res* 30:1359-1364, 2016.
7. Darrall-Jones JD, Jones B, and Till K. Anthropometric and Physical Profiles of English Academy Rugby Union Players. *J Strength Cond Res* 29: 2086-2096, 2015.
8. Enoka RM. Neural adaptations with chronic physical activity. *J Biomech* 30: 447-455, 1997.
9. Faigenbaum AD, McFarland JE, Herman RE, Naclerio F, Ratamess NA, Kang J, and Myer GD. Reliability of the one-repetition-maximum power clean test in adolescent athletes. *J Strength Cond Res* 26: 432-437, 2012.
10. Faigenbaum AD and Myer GD. Resistance training among young athletes: safety, efficacy and injury prevention effects. *Brit J Sports Med* 44: 56-63, 2010.
11. Gabriel DA, Kamen G, and Frost G. Neural adaptations to resistive exercise. *Sports Med* 36: 133-149, 2006.
12. Harries SK, Lubans DR, and Callister R. Comparison of resistance training progression models on maximal strength in sub-elite adolescent rugby union players. *J Sci Med Sport* 19: 163-169, 2016.
13. Haugen T and Buchheit M. Sprint Running Performance Monitoring: Methodological and Practical Considerations. *Sports Med*: 1-16, 2015.
14. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 30: 1-15, 2000.
15. Hopkins WG. Reliability from consecutive pairs of trials (Excel spreadsheet). *Sportsci* 19: 36-42, 2015.
16. Hopkins WG. A spreadsheet for combining outcomes from several subject groups. *Sportsci* 10: 50-53, 2006.
17. Hrysomallis C. Upper-body strength and power changes during a football season. *J Strength Cond Res* 24: 557-559, 2010.
18. Jones B, Emmonds S, Hind K, Nicholson G, Rutherford Z, and Till K. Physical Qualities of International Female Rugby League Players by Playing Position. *J Strength Cond Res* 30: 1333-1340, 2016.
19. Jones B, Till K, Manley A, and McGuigan M. A Multidisciplinary Approach to the Profiling and Interpretation of Fitness Testing Data. *J Aus Strength Cond* In press, 2016.

20. McLaren SJ, Weston M, Smith A, Cramb R, and Portas MD. Variability of physical performance and player match loads in professional rugby union. *J Sci Med Sport* 19: 493-497, 2015.
21. McMaster DT, Gill N, Cronin J, and McGuigan M. The development, retention and decay rates of strength and power in elite rugby union, rugby league and American football: a systematic review. *Sports Med* 43: 367-384, 2013.
22. Mihalik JP, Bell DR, Marshall SW, and Guskiewicz KM. Measurement of head impacts in collegiate football players: an investigation of positional and event- type differences. *Neurosurgery* 61: 1229-1235, 2007.
23. Pyne D. Interpreting the results of fitness testing. In: Proceedings of the International Science and Football Symposium. Melbourne, Australia, 2003.
24. Rhea MR. Determining the magnitude of treatment effects in strength training research through the use of the effect size. *J Strength Cond Res* 18: 918-920, 2004.
25. Ritti-Dias RM, Avelar A, Salvador EP, and Cyrino ES. Influence of previous experience on resistance training on reliability of one-repetition maximum test. *J Strength Cond Res* 25: 1418-1422, 2011.
26. Roe G, Darrall-Jones J, Till K, Phibbs P, Read D, Weakley J, and Jones B. Between-Day Reliability and Sensitivity of Common Fatigue Measures in Rugby Players. *Int J Sports Physiol Perform* 11: 581-586, 2016.
27. Smart DJ, Hopkins WG, and Gill ND. Differences and changes in the physical characteristics of professional and amateur rugby union players. *J Strength Cond Res* 27: 3033-3044, 2013.
28. Smart DJ and Gill ND. Effects of an off-season conditioning program on the physical characteristics of adolescent rugby union players. *J Strength Cond Res* 27: 708-717, 2013.
29. Suchomel TJ, Nimphius S, and Stone MH. The Importance of Muscular Strength in Athletic Performance. *Sports Med* 46:1419-49, 2016.
30. Till K, Darrall-Jones J, Roe G, Weakley J, and Jones B. The influence of training age on the annual development of physical qualities within academy rugby league players. *J Strength Cond Res* (ahead of press) DOI: 10.1519/JSC.0000000000001546.
31. Till K, Jones B, Darrall-Jones J, Emmonds S, and Cooke C. Longitudinal development of anthropometric and physical characteristics within academy rugby league players. *J Strength Cond Res* 29: 1713-1722, 2015.
32. Till K, Jones B, Emmonds S, Tester E, Fahey J, and Cooke C. Seasonal changes in anthropometric and physical characteristics within English academy rugby league players. *J Strength Cond Res* 28: 2689-2696, 2014.
33. Weakley J, Till K, Roe G, Darrall-Jones J, Phibbs P, Read D, and Jones B. Strength and Conditioning Practices in Adolescent Rugby Players: Relationship with Changes in Physical Qualities. *J Strength Cond Res* Published ahead of print, 2017.

Table

Table 1. Measures of reliability for experienced group testing of front squat, chin up, and bench press.

	Squat	Chin Up	Bench Press
Session 1	103.00 ± 17.4	103.70 ± 14.7	92.10 ± 16.50
Session 2	105.50 ± 18.0	105.50 ± 15.40	91.00 ± 15.60
Difference [90% CI]	2.50 [0.45; 4.55]	1.80 [0.24; 2.67]	-1.10 [-3.61; 1.30]
TE (Kg) [90% CI]	2.50 [1.82; 4.11]	1.81 [1.38; 2.69]	3.51 [2.65; 5.31]
CV% [90% CI]	2.90 [2.10; 4.80]	1.70 [1.30; 2.60]	4.50 [3.40; 7.00]
SWC (Kg) (%)	3.39 (3.53)	2.88 (3.01)	3.51 (3.29)
Test Rating	Good	Good	Poor

*TE = typical error of measurement; 90% CI = 90% confidence intervals; TE CV% = typical error of measurement expressed as a coefficient of variation; SWC = smallest worthwhile change

Table

Table 2. Measures of reliability for inexperienced group testing of front squat, chin up, and bench press

	Squat	Chin Up	Bench Press
Session 1	87.5 ± 12.80	95.00 ± 13.03	73.20 ± 15.70
Session 2	87.5 ± 12.60	96.30 ± 12.98	73.40 ± 14.90
Difference [90% CI]	0.00 [-1.37; 1.37]	1.30 [0.00; 2.60]	0.20 [-1.10; 1.50]
TE (Kg) [90% CI]	1.67 [1.22; 2.74]	1.73 [1.26; 2.84]	1.67 [1.23; 2.66]
CV% [90% CI]	1.90 [1.40; 3.10]	1.90 [1.40; 3.10]	2.40 [1.80; 3.80]
SWC (Kg) (%)	2.57 (2.94)	2.73 (2.86)	3.03 (4.17)
Test Rating	Good	Good	Good

*TE = typical error of measurement; 90% CI = 90% confidence intervals; TE CV% = typical error of measurement expressed as a coefficient of variation; SWC = smallest worthwhile change

Table

Table 3. Measures of reliability for combined groups testing of front squat, chin up, and bench press.

	Front squat	Chin Up	Bench press
TE (Kg) [90% CI]	2.26 [1.79; 3.09]	1.74 [1.41; 2.30]	2.81 [2.27; 3.72]
CV% [90% CI]	2.50 [2.00; 3.50]	1.80 [1.40; 2.30]	3.70 [3.00; 4.90]

*TE = typical error of measurement; 90% CI = 90% confidence intervals; TE CV% = typical error of measurement expressed as a coefficient of variation.